

## RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/520,341  
Source: PCT  
Date Processed by STIC: 1-13-05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/10/520,341

TIME: 17:02:32

Input Set : D:\421-67-PCT.ST25.txt

Output Set: N:\CRF4\01132005\J520341.raw

```

3 <110> APPLICANT: The University of North Carolina at Chapel Hill
4     Liu, Jian
6 <120> TITLE OF INVENTION: PURIFIED AND ISOLATED HEPARAN SULFATE 3-O-SULFOTRANSFERASE
7     ISOFORM 5 NUCLEIC ACIDS AND POLYPEPTIDES AND THERAPEUTIC AND
8     SCREENING METHODS USING SAME
10 <130> FILE REFERENCE: 421/67/2
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/520,341
C--> 12 <141> CURRENT FILING DATE: 2005-01-05
12 <150> PRIOR APPLICATION NUMBER: US 60/394,199
13 <151> PRIOR FILING DATE: 2002-07-05
15 <160> NUMBER OF SEQ ID NOS: 12
17 <170> SOFTWARE: PatentIn version 3.2
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 1041
21 <212> TYPE: DNA
22 <213> ORGANISM: Homo sapiens
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (1)..(1041)
29 <400> SEQUENCE: 1
30 atg cta ttc aaa cag cag gcg tgg ctg aga cag aag ctc ctg gtg ctg      48
31 Met Leu Phe Lys Gln Gln Ala Trp Leu Arg Gln Lys Leu Leu Val Leu
32 1          5          10          15
34 gga agc ctt gcc gtt ggg agt ctc ctg tat cta gtc gcc aga gtt ggg      96
35 Gly Ser Leu Ala Val Gly Ser Leu Leu Tyr Leu Val Ala Arg Val Gly
36          20          25          30
38 agc ttg gat agg cta caa ccc att tgc ccc att gaa ggt cga ctg ggt      144
39 Ser Leu Asp Arg Leu Gln Pro Ile Cys Pro Ile Glu Gly Arg Leu Gly
40          35          40          45
42 gga gcc cgc act cag gct gaa ttc cca ctt cgc gcc ctg cag ttt aag      192
43 Gly Ala Arg Thr Gln Ala Glu Phe Pro Leu Arg Ala Leu Gln Phe Lys
44          50          55          60
46 cgt ggc ctg ctg cac gag ttc cgg aag ggc aac gct tcc aag gag cag      240
47 Arg Gly Leu Leu His Glu Phe Arg Lys Gly Asn Ala Ser Lys Glu Gln
48 65          70          75          80
50 gtt cgc ctc cat gac ctg gtc cag cag ctc ccc aag gcc att atc att      288
51 Val Arg Leu His Asp Leu Val Gln Gln Leu Pro Lys Ala Ile Ile Ile
52          85          90          95
54 ggg gtg agg aaa gga ggc aca agg gcc ctg ctt gaa atg ctg aac cta      336
55 Gly Val Arg Lys Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Asn Leu
56          100         105         110
58 cat ccg gca gta gtc aaa gcc tct caa gaa atc cac ttt ttt gat aat      384
59 His Pro Ala Val Val Lys Ala Ser Gln Glu Ile His Phe Phe Asp Asn

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/10/520,341

TIME: 17:02:32

Input Set : D:\421-67-PCT.ST25.txt

Output Set: N:\CRF4\01132005\J520341.raw

```

60          115          120          125
62 gat gag aat tat ggt aag ggc att gag tgg tat agg aaa aag atg cct      432
63 Asp Glu Asn Tyr Gly Lys Gly Ile Glu Trp Tyr Arg Lys Lys Met Pro
64          130          135          140
66 ttt tcc tac cct cag caa atc aca att gaa aag agc cca gca tat ttt      480
67 Phe Ser Tyr Pro Gln Gln Ile Thr Ile Glu Lys Ser Pro Ala Tyr Phe
68 145          150          155          160
70 atc aca gag gag gtt cca gaa agg att tac aaa atg aac tca tcc atc      528
71 Ile Thr Glu Glu Val Pro Glu Arg Ile Tyr Lys Met Asn Ser Ser Ile
72          165          170          175
74 aag ttg ttg atc att gtc agg gag cca acc aca aga gct att tct gat      576
75 Lys Leu Leu Ile Ile Val Arg Glu Pro Thr Thr Arg Ala Ile Ser Asp
76          180          185          190
78 tat act cag gtg cta gag ggg aag gag agg aag aac aaa act tat tac      624
79 Tyr Thr Gln Val Leu Glu Gly Lys Glu Arg Lys Asn Lys Thr Tyr Tyr
80          195          200          205
82 aag ttt gag aag ctg gcc ata gac cct aat aca tgc gaa gtg aac aca      672
83 Lys Phe Glu Lys Leu Ala Ile Asp Pro Asn Thr Cys Glu Val Asn Thr
84          210          215          220
86 aaa tac aaa gca gta aga acc agc atc tac acc aaa cat ctg gaa agg      720
87 Lys Tyr Lys Ala Val Arg Thr Ser Ile Tyr Thr Lys His Leu Glu Arg
88 225          230          235          240
90 tgg ttg aaa tac ttt cca att gag caa ttt cat gtc gtc gat gga gat      768
91 Trp Leu Lys Tyr Phe Pro Ile Glu Gln Phe His Val Val Asp Gly Asp
92          245          250          255
94 cgc ctc atc acg gaa cct ctg cca gaa ctt cag ctc gtg gag aag ttc      816
95 Arg Leu Ile Thr Glu Pro Leu Pro Glu Leu Gln Leu Val Glu Lys Phe
96          260          265          270
98 cta aat ctg cct cca agg ata agt caa tac aat tta tac ttc aat gct      864
99 Leu Asn Leu Pro Pro Arg Ile Ser Gln Tyr Asn Leu Tyr Phe Asn Ala
100          275          280          285
102 acc aga ggg ttt tac tgc ttg cgg ttt aat att atc ttt aat aag tgc      912
103 Thr Arg Gly Phe Tyr Cys Leu Arg Phe Asn Ile Ile Phe Asn Lys Cys
104          290          295          300
106 ctg gcg ggc agc aag ggg cgc att cat cca gag gtg gac ccc tct gtc      960
107 Leu Ala Gly Ser Lys Gly Arg Ile His Pro Glu Val Asp Pro Ser Val
108 305          310          315          320
110 att act aaa ttg cgc aaa ttc ttt cat cct ttt aat caa aaa ttt tac      1008
111 Ile Thr Lys Leu Arg Lys Phe Phe His Pro Phe Asn Gln Lys Phe Tyr
112          325          330          335
114 cag atc act ggg agg aca ttg aac tgg ccc taa      1041
115 Gln Ile Thr Gly Arg Thr Leu Asn Trp Pro
116          340          345
119 <210> SEQ ID NO: 2
120 <211> LENGTH: 346
121 <212> TYPE: PRT
122 <213> ORGANISM: Homo sapiens
124 <400> SEQUENCE: 2
126 Met Leu Phe Lys Gln Gln Ala Trp Leu Arg Gln Lys Leu Leu Val Leu

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/10/520,341

TIME: 17:02:32

Input Set : D:\421-67-PCT.ST25.txt

Output Set: N:\CRF4\01132005\J520341.raw

```

127 1          5          10          15
130 Gly Ser Leu Ala Val Gly Ser Leu Leu Tyr Leu Val Ala Arg Val Gly
131          20          25          30
134 Ser Leu Asp Arg Leu Gln Pro Ile Cys Pro Ile Glu Gly Arg Leu Gly
135          35          40          45
138 Gly Ala Arg Thr Gln Ala Glu Phe Pro Leu Arg Ala Leu Gln Phe Lys
139          50          55          60
142 Arg Gly Leu Leu His Glu Phe Arg Lys Gly Asn Ala Ser Lys Glu Gln
143 65          70          75          80
146 Val Arg Leu His Asp Leu Val Gln Gln Leu Pro Lys Ala Ile Ile Ile
147          85          90          95
150 Gly Val Arg Lys Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Asn Leu
151          100          105          110
154 His Pro Ala Val Val Lys Ala Ser Gln Glu Ile His Phe Phe Asp Asn
155          115          120          125
158 Asp Glu Asn Tyr Gly Lys Gly Ile Glu Trp Tyr Arg Lys Lys Met Pro
159          130          135          140
162 Phe Ser Tyr Pro Gln Gln Ile Thr Ile Glu Lys Ser Pro Ala Tyr Phe
163 145          150          155          160
166 Ile Thr Glu Glu Val Pro Glu Arg Ile Tyr Lys Met Asn Ser Ser Ile
167          165          170          175
170 Lys Leu Leu Ile Ile Val Arg Glu Pro Thr Thr Arg Ala Ile Ser Asp
171          180          185          190
174 Tyr Thr Gln Val Leu Glu Gly Lys Glu Arg Lys Asn Lys Thr Tyr Tyr
175          195          200          205
178 Lys Phe Glu Lys Leu Ala Ile Asp Pro Asn Thr Cys Glu Val Asn Thr
179          210          215          220
182 Lys Tyr Lys Ala Val Arg Thr Ser Ile Tyr Thr Lys His Leu Glu Arg
183 225          230          235          240
186 Trp Leu Lys Tyr Phe Pro Ile Glu Gln Phe His Val Val Asp Gly Asp
187          245          250          255
190 Arg Leu Ile Thr Glu Pro Leu Pro Glu Leu Gln Leu Val Glu Lys Phe
191          260          265          270
194 Leu Asn Leu Pro Pro Arg Ile Ser Gln Tyr Asn Leu Tyr Phe Asn Ala
195          275          280          285
198 Thr Arg Gly Phe Tyr Cys Leu Arg Phe Asn Ile Ile Phe Asn Lys Cys
199          290          295          300
202 Leu Ala Gly Ser Lys Gly Arg Ile His Pro Glu Val Asp Pro Ser Val
203 305          310          315          320
206 Ile Thr Lys Leu Arg Lys Phe Phe His Pro Phe Asn Gln Lys Phe Tyr
207          325          330          335
210 Gln Ile Thr Gly Arg Thr Leu Asn Trp Pro
211          340          345
214 <210> SEQ ID NO: 3
215 <211> LENGTH: 307
216 <212> TYPE: PRT
217 <213> ORGANISM: Homo sapiens
219 <400> SEQUENCE: 3
221 Met Ala Ala Leu Leu Leu Gly Ala Val Leu Leu Val Ala Gln Pro Gln

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/10/520,341

TIME: 17:02:32

Input Set : D:\421-67-PCT.ST25.txt

Output Set: N:\CRF4\01132005\J520341.raw

```

222 1          5          10          15
225 Leu Val Pro Ser Arg Pro Ala Glu Leu Gly Gln Gln Glu Leu Leu Arg
226          20          25          30
229 Lys Ala Gly Thr Leu Gln Asp Asp Val Arg Asp Gly Val Ala Pro Asn
230          35          40          45
233 Gly Ser Ala Gln Gln Leu Pro Gln Thr Ile Ile Ile Gly Val Arg Lys
234          50          55          60
237 Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Ser Leu His Pro Asp Val
238 65          70          75          80
241 Ala Ala Ala Glu Asn Glu Val His Phe Phe Asp Trp Glu Glu His Tyr
242          85          90          95
245 Ser His Gly Leu Gly Trp Tyr Leu Ser Gln Met Pro Phe Ser Trp Pro
246          100         105         110
249 His Gln Leu Thr Val Glu Lys Thr Pro Ala Tyr Phe Thr Ser Pro Lys
250          115         120         125
253 Val Pro Glu Arg Val Tyr Ser Met Asn Pro Ser Ile Arg Leu Leu Leu
254          130         135         140
257 Ile Leu Arg Asp Pro Ser Glu Arg Val Leu Ser Asp Tyr Thr Gln Val
258 145         150         155         160
261 Phe Tyr Asn His Met Gln Lys His Lys Pro Tyr Pro Ser Ile Glu Glu
262          165         170         175
265 Phe Leu Val Arg Asp Gly Arg Leu Asn Val Asp Tyr Lys Ala Leu Asn
266          180         185         190
269 Arg Ser Leu Tyr His Val His Met Gln Asn Trp Leu Arg Phe Phe Pro
270          195         200         205
273 Leu Arg His Ile His Ile Val Asp Gly Asp Arg Leu Ile Arg Asp Pro
274          210         215         220
277 Phe Pro Glu Ile Gln Lys Val Glu Arg Phe Leu Lys Leu Ser Pro Gln
278 225         230         235         240
281 Ile Asn Ala Ser Asn Phe Tyr Phe Asn Lys Thr Lys Gly Phe Tyr Cys
282          245         250         255
285 Leu Arg Asp Ser Gly Arg Asp Arg Cys Leu His Glu Ser Lys Gly Arg
286          260         265         270
289 Ala His Pro Gln Val Asp Pro Lys Leu Leu Asn Lys Leu His Glu Tyr
290          275         280         285
293 Phe His Glu Pro Asn Lys Lys Phe Phe Glu Leu Val Gly Arg Thr Phe
294          290         295         300
297 Asp Trp His
298 305
301 <210> SEQ ID NO: 4
302 <211> LENGTH: 406
303 <212> TYPE: PRT
304 <213> ORGANISM: Homo sapiens
306 <400> SEQUENCE: 4
308 Met Ala Pro Pro Gly Pro Ala Ser Ala Leu Ser Thr Ser Ala Glu Pro
309 1          5          10          15
312 Leu Ser Arg Ser Ile Phe Arg Lys Phe Leu Leu Met Leu Cys Ser Leu
313          20          25          30
316 Leu Thr Ser Leu Tyr Val Phe Tyr Cys Leu Ala Glu Arg Cys Gln Thr

```

## RAW SEQUENCE LISTING

DATE: 01/13/2005

PATENT APPLICATION: US/10/520,341

TIME: 17:02:32

Input Set : D:\421-67-PCT.ST25.txt

Output Set: N:\CRF4\01132005\J520341.raw

```

317          35          40          45
320 Leu Ser Gly Pro Val Val Gly Leu Ser Gly Gly Gly Glu Glu Ala Gly
321          50          55          60
324 Ala Pro Gly Gly Gly Val Leu Ala Gly Gly Pro Arg Glu Leu Ala Val
325 65          70          75          80
328 Trp Pro Ala Ala Ala Gln Arg Lys Arg Leu Leu Gln Leu Pro Gln Trp
329          85          90          95
332 Arg Arg Arg Arg Pro Pro Ala Pro Arg Asp Asp Gly Glu Glu Ala Ala
333          100          105          110
336 Trp Glu Glu Glu Ser Pro Gly Leu Ser Gly Gly Pro Gly Gly Ser Gly
337          115          120          125
340 Ala Gly Ser Thr Val Ala Glu Ala Pro Pro Gly Thr Leu Ala Leu Leu
341          130          135          140
344 Leu Asp Glu Gly Ser Lys Gln Leu Pro Gln Ala Ile Ile Ile Gly Val
345 145          150          155          160
348 Lys Lys Gly Gly Thr Arg Ala Leu Leu Glu Phe Leu Arg Val His Pro
349          165          170          175
352 Asp Val Arg Ala Val Gly Ala Glu Pro His Phe Phe Asp Arg Ser Tyr
353          180          185          190
356 Asp Lys Gly Leu Ala Trp Tyr Arg Asp Leu Met Pro Arg Thr Leu Asp
357          195          200          205
360 Gly Gln Ile Thr Met Glu Lys Thr Pro Ser Tyr Phe Val Thr Arg Glu
361          210          215          220
364 Ala Pro Ala Arg Ile Ser Ala Met Ser Lys Asp Thr Lys Leu Ile Val
365 225          230          235          240
368 Val Val Arg Asp Pro Val Thr Arg Ala Ile Ser Asp Tyr Thr Gln Thr
369          245          250          255
372 Leu Ser Lys Arg Pro Asp Ile Pro Thr Phe Glu Ser Leu Thr Phe Lys
373          260          265          270
376 Asn Arg Thr Ala Gly Leu Ile Asp Thr Ser Trp Ser Ala Ile Gln Ile
377          275          280          285
380 Gly Ile Tyr Ala Lys His Leu Glu His Trp Leu Arg His Phe Pro Ile
381          290          295          300
384 Arg Gln Met Leu Phe Val Ser Gly Glu Arg Leu Ile Ser Asp Pro Ala
385 305          310          315          320
388 Gly Glu Leu Gly Arg Val Gln Asp Phe Leu Gly Leu Lys Arg Ile Ile
389          325          330          335
392 Thr Asp Lys His Phe Tyr Phe Asn Lys Thr Lys Gly Phe Pro Cys Leu
393          340          345          350
396 Lys Lys Ala Glu Gly Ser Ser Arg Pro His Cys Leu Gly Lys Thr Lys
397          355          360          365
400 Gly Arg Thr His Pro Glu Ile Asp Arg Glu Val Val Arg Arg Leu Arg
401          370          375          380
404 Glu Phe Tyr Arg Pro Phe Asn Leu Lys Phe Tyr Gln Met Thr Gly His
405 385          390          395          400
408 Asp Phe Gly Trp Asp Gly
409          405
412 <210> SEQ ID NO: 5
413 <211> LENGTH: 390

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/520,341

DATE: 01/13/2005  
TIME: 17:02:33

Input Set : D:\421-67-PCT.ST25.txt  
Output Set: N:\CRF4\01132005\J520341.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,  
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:6,7,8,9,10,11,12

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/520,341

DATE: 01/13/2005

TIME: 17:02:33

Input Set : D:\421-67-PCT.ST25.txt

Output Set: N:\CRF4\01132005\J520341.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date